Semantic Integration
in the
Information Flow Framework

Robert E. Kent

Ontologos
550 SW Staley Dr, Pullman WA 99163, USA
rekent@ontologos.org

Abstract. The Information Flow Framework (IFF) [1] is a descriptive category metatheory currently under development, which is being offered as the structural aspect of the Standard Upper Ontology (SUO). The architecture of the IFF is composed of metalevels, namespaces and meta-ontologies. The main application of the IFF is institutional: the notion of institutions and their morphisms are being axiomatized in the upper metalevels of the IFF, and the lower metalevel of the IFF has axiomatized various institutions in which semantic integration has a natural expression as the colimit of theories. Some of the ideas used in this paper first appeared in papers by Joseph Goguen [2] and the author [3], and discussions on the SUO email list. See also the companion paper [4].

Keywords: descriptive category metatheory, institutions, semantic integration

"*Philosophy cannot become scientifically healthy without an immense technical vocabulary. We can hardly imagine our great-grandsons turning over the leaves of this dictionary without amusement over the paucity of words with which their grandsires attempted to handle metaphysics and logic. Long before that day, it will have become indispensably requisite, too, that each of these terms should be confined to a single meaning which, however broad, must be free from all vagueness. This will involve a revolution in terminology; for in its present condition a philosophical thought of any precision can seldom be expressed without lengthy explanations.*"
— Charles Sanders Peirce, Collected Papers 8:169

1   The Information Flow Framework

The IEEE P1600.1 Standard Upper Ontology (SUO) project aims to specify an upper ontology that will provide a structure and a set of general concepts upon which object-level ontologies can be constructed. The Information Flow Framework (IFF) [1] is a descriptive category metatheory that represents the structural aspect of the SUO containing meta, generic and abstract concepts. To be independent of the particular logic used in object-level ontologies, the

IFF represents and manipulates ontological structures within the metatheory of institutions. The IFF has work-in-progress axiomatizations for (amongst others) the institutions of information flow (**IF**), equational logic (**EQL**), order-sorted first order logic (**FOL**) and the common logic standard (**CL**), and is developing an axiomatization for the metatheory of institutions itself. Its institutional approach to logical semantics provides a principled framework for the modular design of object-level ontologies; in particular, for the "lattice of theories" approach to ontological organization. A major requirement of the SUO, called the "lattice of theories", is to create a framework which can support an open-ended number of theories (potentially infinite) organized in a lattice together with systematic metalevel techniques for moving from one to another, for testing their adequacy for any given problem, and for mixing, matching, combining, and transforming them to whatever form is appropriate for whatever problem anyone is trying to solve (John Sowa [5] and the SUO archive). Within the IFF representation, the lattice of theories is the fibring or indexing of the context of theories by the context of signatures (aka languages) — all definable within institutions. Semantic integration of ontologies can be represented by (1) aligning ontologies within a diagram of theories and (2) fusing aligned ontologies via the colimit of that diagram of theories.

The IFF has gone through two phases of development, and is now well within its third phase, which involves reorganization of the IFF core hierarchy using the "adjunctive axiomatization" technique as illustrated by Lawvere and Rosebrugh [6]. A fourth phase is envisioned in the future, where the concepts of fibrations and indexed categories will be axiomatized. These are important for the axiomatizations of institutions and fibring logics. During the IFF development, four concepts have eventually emerged as important. In chronological order these are (1) the conceptual warrant principle, (2) the categorical design principle, (3) the institutional logic principle, and (4) the metastack concept.

– **Conceptual Warrant:** *All IFF terminology should require conceptual warrant for their existence: any term that appears in (and is axiomatized by) a metalanguage should reference a concept needed in a lower metalevel or object level axiomatization.*
  The principle of conceptual warrant originated in phase one of the IFF development. Warrant means evidence for or token of authorization. Conceptual warrant is an adaptation of the librarianship notion of literary warrant. According to the Library of Congress, its collections serve as the literary warrant (i.e., the literature on which the controlled vocabulary is based) for the Library of Congress subject headings system. In the same fashion, the object-level and lower metalevel terminology of the IFF serves as the conceptual warrant for the IFF upper metalevel axiomatization.

– **Categorical Design:** *The design of a module should adhere to the property that its axiomatic representation is strictly category-theoretic: All axioms use terms from the metalanguage at that metalevel. No axioms use explicit logical notation: no variables, quantifications or logical connectives are used.*

The principle of categorical design originated in phase two of the IFF development. The goal of this principle is to simplify the IFF axiomatization — first order expression would be reduced to term-rewriting. The peripheral (non-core) modules in the lower IFF metalevel have the tripartite form: outer category namespace, inner object and morphism namespaces. The outer namespace fully conforms to the categorical design principle. The inner namespaces conform to it to a great extent (80–90%).

– **Institutional Logic:** *All logics used in the IFF application should be formulated as institutions.*
  The principle of institutional logic[1] originated in phase three of the IFF development. The theories of information flow and formal concept analysis (and hence, effectively the theory of institutions) have been used throughout the IFF development. This use has centered on the "truth construction", which is institutional (see Subsection 2.2). As noted in [7], truth is not dyadic between models and sentences, but triadic between models, sentences and signatures. This corresponds to the contextual dependency in the semiotics of Charles Sanders Peirce, which uses signs, objects and interpretants. In formal concept analysis, triadic concept lattices [8] have been used to formalize this. However, the representation of contexts as the logics of institutions has greater advantages.

– **Metastack:**
  The idea of the IFF metastack also originated in phase three of the IFF development. The IFF metastack is the IFF core hierarchy. The content of the IFF metastack is the axiomatization for sets, functions and binary relations in four core modules at four different set-theoretic levels, the "small" in the lower metalevel, the "large" in the upper metalevel, the "very large" in the top metalevel and the "generic" in the ur metalevel. This includes an axiomatization for generic categories in the ur metalevel, an axiomatization for very large finite limits in the top metalevel, an axiomatization for large exponents and finite colimits in the upper metalevel, and an axiomatization for small subobject classifiers and general limits/colimits in the lower metalevel. Although, this effectively distributes topos axioms over the four metalevels, there is still an effort to follow conceptual warrant.

The IFF architecture consists of metalevels, namespaces and meta-ontologies. Within each level, the terminology is partitioned into namespaces. The number of namespaces and the content may vary over time: new namespaces may be created or old namespaces may be deprecated, and new terminology and axiomatization within any particular namespace may change (new versions). In addition, within each level, various namespaces are collected together into meaningful composites called meta-ontologies. At any particular metalevel, these meta-ontologies cover all the namespaces at that level, but they may overlap. The number of meta-ontologies and the content of any meta-ontology may vary over time: new meta-

---

[1] Suggested to the author by Joseph Goguen (personal communication).

ontologies may be created or old meta-ontologies may be deprecated, and new namespaces within any particular meta-ontology may change (new versions).

The IFF terminology is managed in terms of namespace prefixes — each namespace is given a unique prefix (with perhaps a few synonyms) in order to avoid clash of terminology. The architecture of the IFF namespace mechanism is flat — namespace prefixes are like tags: by using namespace prefixes the complete IFF terminology is the disjoint union of the terminology in the IFF namespaces. The IFF architecture can be thought of as a two dimensional structure consisting of metalevels, which are partitioned into top-level namespaces representing basic concepts such as "category", "graph" or "institution". The various levels are indexed by the natural numbers or their language correlates, starting with the object level indexed by zero. Overall, various namespaces may have the same name, since they represent the same basic concept at different metalevels. To locate any namespace one can use its level-concept pair. For example, the namespace that axiomatizes large categories is located at the "second" metalevel and represents the "category" concept. It is assumed that each basic concept has a particular metalevel that is in common use. For such namespaces, the level notation need not be used.

There are thousands of terms in the IFF. Terms divide into two classes, which we can call "usable terms" and "supporting terms". An IFF term, which is defined in a particular namespace on a particular metalevel, is a *usable* IFF term when it is used by at least one other term in another namespace on that metalevel or on a level below that one. An IFF term, which is defined in a particular namespace on a particular metalevel, is a *supporting* IFF term when it is used by another term in that same namespace. Because of conceptual warrant, all IFF terms should be usable or supporting, and perhaps both. Hence, all IFF terms are necessary, but most IFF terms are "conceptually derived". This means that they are a conceptual composite of two or more basic IFF terms. An IFF term is a *basic* IFF term when it is not the conceptual composite of two or more other IFF terms. Currently, there are about one hundred basic IFF terms.

There are four IFF metalevels: lower, upper, top and ur. Each metalevel services the level below: the ur metalevel services the top metalevel, the ur and top metalevels service the upper metalevel, the ur, top and upper metalevels service the lower metalevel, and the ur, top, upper and lower metalevels service the object-level. There is one metalanguage associated with each metalevel. That is, each metalevel has an associated metalanguage, whose *old* terminology is the terminology of the metalanguage associated with the metalevel immediately above and whose *new* terminology is defined by the various meta-ontologies at that metalevel. Any metalanguage can be used by the meta-ontologies and ontologies at all lower levels. There results a hierarchy of IFF metalanguages `meta-lower` $\sqsupseteq$ `meta-upper` $\sqsupseteq$ `meta-top` $\sqsupseteq$ `meta-ur` $\sqsupseteq$ `metashell` coordinated with the IFF metastack, where the logical shell called `metashell` enables a lisp-like first-order expression using connectives and quantifiers with a restricted quantification format. The `meta-ur` metalanguage is special — it axiomatizes the IFF metastack.

2   Semantic Integration with Institutions

Abstract semantic integration can be defined and axiomatized in the metatheory of institutions, where it has both an intrinsic and an extrinsic formulation. In this paper, we discuss the intrinsic formulation in terms of the colimits of theories. A later paper will discuss the extrinsic formulation.

2.1   Institutions

An institution $\Im = \langle \mathsf{Sign}_\Im, \mathsf{mod}_\Im, \mathsf{sen}_\Im, \models_\Im \rangle$ [9] has an abstract category $\mathsf{Sign}_\Im$ of signatures Σ, a sentence fiber functor $\mathsf{sen}_\Im : \mathsf{Sign}_\Im \to \mathsf{Set}$ indexing abstract sentences $\mathsf{sen}_\Im(\Sigma)$ by signatures Σ, a model fiber (reduct) functor $\mathsf{mod}_\Im : \mathsf{Sign}_\Im \to \mathsf{CAT}^{\mathsf{op}}$ indexing abstract models $\mathsf{mod}_\Im(\Sigma)$ by signatures Σ, and a function $\models_\Im : |\mathsf{Sign}_\Im| \to \mathsf{REL}$ indexing abstract satisfaction relations $\models_{\Im,\Sigma} \subseteq |\mathsf{mod}|_\Im(\Sigma) \times \mathsf{sen}_\Im(\Sigma)$ by signatures Σ. An institution must satisfy the *satisfaction condition*: $|\mathsf{mod}|_\Im(\sigma)(m_2) \models_{\Im,\Sigma_1} s_1 \underline{\text{iff}} \ m_2 \models_{\Im,\Sigma_2} \mathsf{sen}_\Im(\sigma)(s_1)$, for any signature morphism $\sigma : \Sigma_1 \to \Sigma_2$, any target model $m_2 \in |\mathsf{mod}|_\Im(\Sigma_2)$ and any source sentence $s_1 \in \mathsf{sen}_\Im(\Sigma_1)$. The satisfaction condition expresses the invariance of truth under change of notation. Satisfaction does not use model morphisms, morphisms in $\mathsf{mod}_\Im(\Sigma)$, and hence is expressed in terms of the underlying model functor $|\mathsf{mod}|_\Im = \mathsf{mod}_\Im \circ |-|^{\mathsf{op}} : \mathsf{Sign}_\Im \to \mathsf{SET}^{\mathsf{op}}$. Examples of institutions include: first order logic with first order structures as models, many sorted equational logic with abstract algebras as models, Horn clause logic, and variants of higher order and of modal logic.

For every signature Σ of $\Im$, there is a complete preorder $\mathsf{th}^{\vdash}_\Im(\Sigma)$, whose elements, called Σ-theories of $\Im$, are subsets of sentences, whose order is entailment order, whose join operator is intersection, and whose meet operator is union. Elements of theories are called axioms. We can extend satisfaction to theories: a Σ-model m satisfies (is a model of) a Σ-theory T, denoted $m \models_{\Im,\Sigma} T$, when m satisfies each axiom $s \in T$; that is, when $m \models_{\Im,\Sigma} s$ for all axioms $s \in T$. A theory $T \in \mathsf{th}^{\vdash}_\Im(\Sigma)$ entails a sentence $s \in \mathsf{sen}_\Im(\Sigma)$, denoted by $T \vdash_{\Im,\Sigma} s$, when $m \models_{\Im,\Sigma} T$ implies $m \models_{\Im,\Sigma} s$ for any model m. Such a sentence is called a theorem of T. The set of all theorems of T, called the closure of T, is denoted by $\mathsf{clo}_{\Im,\Sigma}(T) = T^\bullet$. Closure is monotonic, increasing and idempotent. If $T_1 \subseteq T_2$ then $T_1^\bullet \subseteq T_2^\bullet$. Any axiom is a theorem: $T \subseteq T^\bullet$. The closure of the closure is the closure: $T^{\bullet\bullet} = T^\bullet$. A Σ-theory T_1 entails a Σ-theory T_2 when T_1 entails every axiom of T_2; that is, when $T_1^\bullet \supseteq T_2$. Then we also say that T_2 is a generalization of T_1 or that T_1 is a specialization of a T_2, and denote this entailment order by $T_1 \vdash_{\Im,\Sigma} T_2$. Entailment is equivalent to closure-subset: $T_1 \vdash_{\Im,\Sigma} T_2$ iff $T_1^\bullet \supseteq T_2^\bullet$. Hence, two Σ-theories are entailment equivalent $T_1 \equiv T_2$ iff they have the same closure $T_1^\bullet = T_2^\bullet$. In particular, any theory T is equivalent to its closure in entailment preorder: $T \equiv T^\bullet$. A Σ-theory is closed when it is equal to, not just equivalent to, its closure: $T = T^\bullet$. Closed theories $T^\bullet$ are in one-one correspondence with equivalence classes $[T]$; that is, closed theories form a set of representative elements for the entailment equivalence relation. Let $\mathsf{th}^\bullet_\Im(\Sigma)$ denote the complete lattice, whose elements are the closed Σ-theories of $\Im$, whose

order is reverse subset inclusion, whose join operator is intersection, and whose meet operator is union followed by closure.

Every model $m \in |\mathsf{mod}|_{\Im}(\Sigma)$ has an associated (closed) theory $\iota_{\Im}^{\bullet}(\Sigma)(m) = \{s \in \mathsf{sen}_{\Im}(\Sigma) \mid m \models_{\Im,\Sigma} s\}$ which is maximal in subset order (minimal in entailment order) for all theories that m satisfies. This definition gives a model closed embedding function $\iota_{\Im}^{\bullet}(\Sigma) : |\mathsf{mod}|_{\Im}(\Sigma) \to \mathsf{th}_{\Im}^{\bullet}(\Sigma)$ and a model entailment embedding function $\mathsf{mod\text{-}th}_{\Im}(\Sigma) = \iota_{\Im}^{\bullet}(\Sigma) \cdot \mathsf{incl}_{\Im,\Sigma} : |\mathsf{mod}|_{\Im}(\Sigma) \to \mathsf{th}_{\Im}^{\vdash}(\Sigma)$. Every sentence $s \in \mathsf{sen}_{\Im}(\Sigma)$ has an associated (closed) theory $\mathsf{sen\text{-}th}_{\Im}(\Sigma)(s) = \{s' \in \mathsf{sen}_{\Im}(\Sigma) \mid \{s\} \vdash_{\Im,\Sigma} s'\}$. This definition gives a sentence entailment embedding function $\mathsf{sen\text{-}th}_{\Im}(\Sigma) : \mathsf{sen}_{\Im}(\Sigma) \to \mathsf{th}_{\Im}^{\vdash}(\Sigma)$ and a sentence closed embedding function $\tau_{\Im}^{\bullet}(\Sigma) = \mathsf{sen\text{-}th}_{\Im}(\Sigma) \cdot \mathsf{clo}_{\Im,\Sigma} : \mathsf{sen}_{\Im}(\Sigma) \to \mathsf{th}_{\Im}^{\bullet}(\Sigma)$, since closure is idempotent.

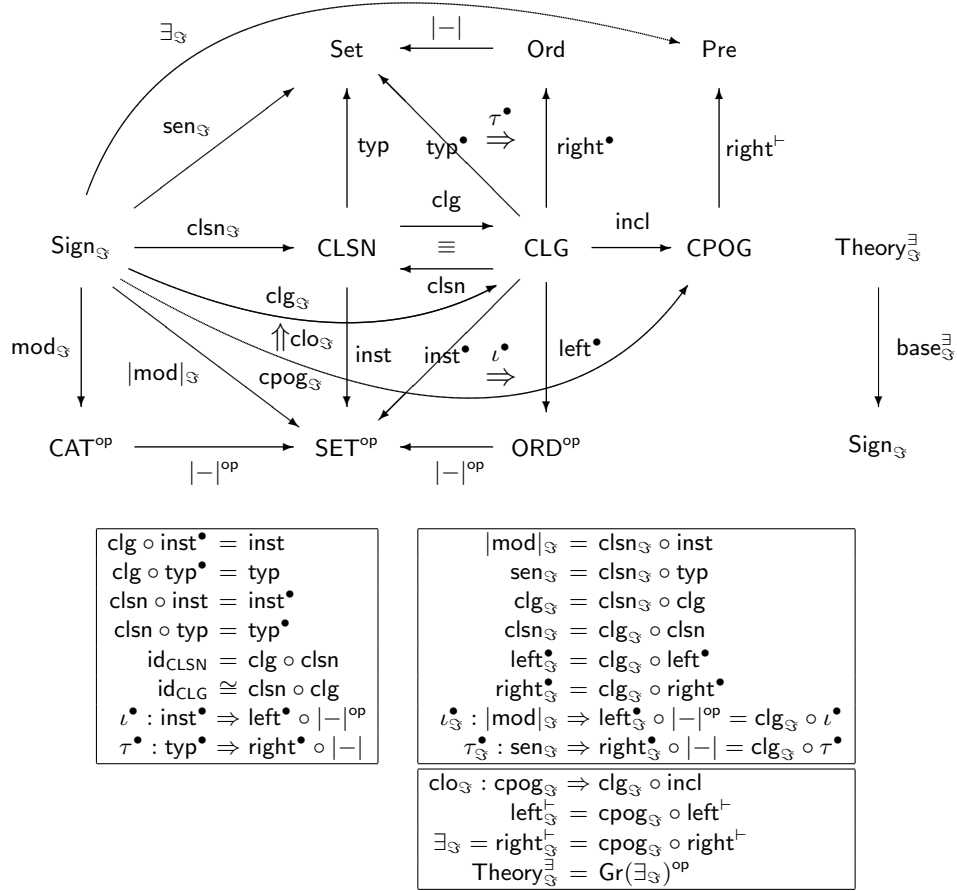**Fig. 1.** The Architecture of Institutions

2.2   Truth Construction

The components of an institution $\Im$, which are illustrated on the left side of Figure 1, can be packed together as a classification functor $\mathsf{clsn}_\Im : \mathsf{Sign}_\Im \to \mathsf{CLSN}$ into the category of (large) classifications and infomorphisms [10], where for every signature Σ the satisfaction relation forms the classification $\mathsf{clsn}_\Im(\Sigma) = \langle |\mathsf{mod}|_\Im(\Sigma), \mathsf{sen}_\Im(\Sigma), \models_{\Im,\Sigma} \rangle$ and for every signature morphism $\sigma : \Sigma_1 \to \Sigma_2$ the satisfaction condition states the fundamental condition for the infomorphism $\mathsf{clsn}_\Im(\sigma) = \langle |\mathsf{mod}|_\Im(\sigma), \mathsf{sen}_\Im(\sigma) \rangle : \mathsf{clsn}_\Im(\Sigma_1) \rightleftharpoons \mathsf{clsn}_\Im(\Sigma_2)$. However, the category of classifications CLSN is equivalent to the category of (large) concept lattices CLG (complete lattices with two-sided generators) and concept morphisms: $\mathsf{CLSN} \equiv \mathsf{CLG}$ [11], an equivalence mediated by the concept lattice functor $\mathsf{clg} : \mathsf{CLSN} \to \mathsf{CLG}$ and the classification functor $\mathsf{clsn} : \mathsf{CLG} \to \mathsf{CLSN}$. Several isomorphic representations of concept lattices are possible — a full two-sided version with both extents and intents, an extent-only version and an intent-only version. In institution theory, we normally choose the intent-only version, where $\mathsf{clg}_\Im(\Sigma)$ has the complete lattice component $\mathsf{th}^\bullet_\Im(\Sigma)$, whose elements (formal concepts) are the closed Σ-theories of $\Im$.

Hence, an alternate expression for an institution is a concept lattice functor $\mathsf{clg}_\Im : \mathsf{Sign}_\Im \to \mathsf{CLG}$, where for every signature Σ the associated concept lattice is the quintuple $\mathsf{clg}_\Im(\Sigma) = \langle |\mathsf{mod}|_\Im(\Sigma), \iota^\bullet_\Im(\Sigma), \mathsf{th}^\bullet_\Im(\Sigma), \tau^\bullet_\Im(\Sigma), \mathsf{sen}_\Im(\Sigma) \rangle$ and for every signature morphism $\sigma : \Sigma_1 \to \Sigma_2$ the associated concept morphism is the quadruple $\mathsf{clg}_\Im(\sigma) = \langle |\mathsf{mod}|_\Im(\sigma), \mathsf{left}^\bullet_\Im(\sigma), \mathsf{right}^\bullet_\Im(\sigma), \mathsf{sen}_\Im(\sigma) \rangle : \mathsf{clg}_\Im(\Sigma_1) \rightleftharpoons \mathsf{clg}_\Im(\Sigma_2)$. The left closed operator $\mathsf{left}^\bullet_\Im(\sigma) : \mathsf{th}^\bullet_\Im(\Sigma_1) \leftarrow \mathsf{th}^\bullet_\Im(\Sigma_2)$, which is a join-preserving monotonic function, is the substitution operator $\mathsf{subst}_\Im(\sigma) = \mathsf{sen}_\Im(\sigma)^{-1}$, restricted to closed theories. The right closed operator $\mathsf{right}^\bullet_\Im(\sigma) = \exists_\Im(\sigma) \cdot \mathsf{clo}_\Im(\Sigma_2) : \mathsf{th}^\bullet_\Im(\Sigma_1) \to \mathsf{th}^\bullet_\Im(\Sigma_2)$, which is a meet-preserving monotonic function, is the existential quantification operator, restricted to closed theories, composed with the target closure operator. The left-right pair of operators forms a Galois connection $\phi^\bullet_\Im(\sigma) = \langle \mathsf{left}^\bullet_\Im(\sigma), \mathsf{right}^\bullet_\Im(\sigma) \rangle : \mathsf{th}^\bullet_\Im(\Sigma_2) \rightleftharpoons \mathsf{th}^\bullet_\Im(\Sigma_1)$. The model operator is compatible with the left closed operator (via model-embedding): $|\mathsf{mod}|_\Im(\sigma) \cdot \iota^\bullet_\Im(\Sigma_1) = \iota^\bullet_\Im(\Sigma_2) \cdot \mathsf{left}^\bullet_\Im(\sigma)$, and the sentence operator is compatible with the right closed operator (via sentence-embedding): $\tau^\bullet_\Im(\Sigma_1) \cdot \mathsf{right}^\bullet_\Im(\sigma) = \mathsf{sen}_\Im(\sigma) \cdot \tau^\bullet_\Im(\Sigma_2)$. The equivalence of the two functors, $\mathsf{clsn}_\Im$ and $\mathsf{clg}_\Im$, is expressed by the identities: $\mathsf{clg}_\Im = \mathsf{clsn}_\Im \circ \mathsf{clg}$ and $\mathsf{clsn}_\Im = \mathsf{clg}_\Im \circ \mathsf{clsn}$.

The category CPOG of (large) complete preorders with two-sided generators is adjointly related to the category CLG of concept lattices: $\mathsf{CLG} \dashv\vdash \mathsf{CPOG}$, an adjointness mediated by the inclusion functor $\mathsf{incl} : \mathsf{CLG} \to \mathsf{CPOG}$ and the equivalence quotient functor $\mathsf{quo} : \mathsf{CPOG} \to \mathsf{CLG}$. In CPOG, the orders do not necessarily satisfy antisymmetry — a pair of elements may be equivalent but not equal. An institution $\Im$ has an associated complete preorder functor $\mathsf{cpog}_\Im : \mathsf{Sign}_\Im \to \mathsf{CPOG}$, which is compositionally adjoint to the concept lattice functor (and hence, also to the classification functor). For every signature Σ, the associated complete preorder is the quintuple $\mathsf{cpog}_\Im(\Sigma) = \langle |\mathsf{mod}|_\Im(\Sigma), \mathsf{mod\text{-}th}_\Im(\Sigma), \mathsf{th}^\vdash_\Im(\Sigma), \mathsf{sen\text{-}th}_\Im(\Sigma), \mathsf{sen}_\Im(\Sigma) \rangle$ and for every signature morphism $\sigma : \Sigma_1 \to \Sigma_2$ the associated complete preorder mor-

phism is the quadruple $\mathsf{cpog}_\Im(\sigma) = \langle |\mathsf{mod}|_\Im(\sigma), \mathsf{left}_\Im^\vdash(\sigma), \mathsf{right}_\Im^\vdash(\sigma), \mathsf{sen}_\Im(\sigma)\rangle$: $\mathsf{cpog}_\Im(\Sigma_1) \rightleftharpoons \mathsf{cpog}_\Im(\Sigma_2)$. For every signature morphism $\sigma : \Sigma_1 \rightarrow \Sigma_2$, the existential quantification monotonic function $\exists_\Im(\sigma) : \mathsf{th}_\Im^\vdash(\Sigma_1) \rightarrow \mathsf{th}_\Im^\vdash(\Sigma_2)$ maps a source theory $T_1 \in \mathsf{th}_\Im(\Sigma_1)$ to the target theory $\exists_\Im(\sigma)(T_1) = \{s_2 \in \mathsf{sen}_\Im(\Sigma_2) \mid \exists_{s_1 \in \mathsf{sen}_\Im(\Sigma_1)} (\mathsf{sen}_\Im(\sigma)(s_1) = s_2) \ \& \ s_1 \in T_1\} \in \mathsf{th}_\Im(\Sigma_2)$, and the substitution monotonic function is the inverse image of the sentence function $\mathsf{subst}_\Im(\sigma) = \mathsf{sen}_\Im(\sigma)^{-1} : \mathsf{th}_\Im^\vdash(\Sigma_1) \leftarrow \mathsf{th}_\Im^\vdash(\Sigma_2)$. The left entailment operator $\mathsf{left}_\Im^\vdash(\sigma) = \mathsf{clo}_\Im(\Sigma_2) \cdot \mathsf{subst}_\Im(\sigma) : \mathsf{th}_\Im^\vdash(\Sigma_1) \leftarrow \mathsf{th}_\Im^\vdash(\Sigma_2)$, which is a join-preserving monotonic function, is the composition of target closure with substitution. The right entailment operator $\mathsf{right}_\Im^\vdash(\sigma) = \exists_\Im(\sigma) : \mathsf{th}_\Im^\vdash(\Sigma_1) \rightarrow \mathsf{th}_\Im^\vdash(\Sigma_2)$, which is a meet-preserving monotonic function, is existential quantification. The left-right pair of operators forms a Galois connection $\phi_\Im^\vdash(\sigma) = \langle \mathsf{left}_\Im^\vdash(\sigma), \mathsf{right}_\Im^\vdash(\sigma)\rangle$: $\mathsf{th}_\Im^\vdash(\Sigma_2) \rightleftharpoons \mathsf{th}_\Im^\vdash(\Sigma_1)$. The model operator is compatible with the left entailment operator (via model-embedding): $|\mathsf{mod}|_\Im(\sigma) \cdot \mathsf{mod\text{-}th}_\Im(\Sigma_1) = \mathsf{mod\text{-}th}_\Im(\Sigma_2) \cdot \mathsf{left}_\Im^\vdash(\sigma)$, and the sentence operator is compatible with the right entailment operator (via sentence-embedding): $\mathsf{sen\text{-}th}_\Im(\Sigma_1) \cdot \mathsf{right}_\Im^\vdash(\sigma) = \mathsf{sen}_\Im(\sigma) \cdot \mathsf{sen\text{-}th}_\Im(\Sigma_2)$.

For every signature Σ, the closure operator $\mathsf{clo}_{\Im,\Sigma} : \mathsf{th}_\Im^\vdash(\Sigma) \rightarrow \mathsf{th}_\Im^\bullet(\Sigma)$ and inclusion $\mathsf{incl}_{\Im,\Sigma} : \mathsf{th}_\Im^\vdash(\Sigma) \leftarrow \mathsf{th}_\Im^\bullet(\Sigma)$ are monotonic, and form a Galois connection $\rho_\Im(\Sigma) = \langle \mathsf{incl}_\Im(\Sigma), \mathsf{clo}_\Im(\Sigma)\rangle : \mathsf{th}_\Im^\bullet(\Sigma) \rightleftharpoons \mathsf{th}_\Im^\vdash(\Sigma)$. The left (entailment, closed) monotonic function commutes with the inclusion monotonic function: $\mathsf{incl}_\Im(\Sigma_2) \cdot \mathsf{left}_\Im^\vdash(\sigma) = \mathsf{left}_\Im^\bullet(\sigma) \cdot \mathsf{incl}_\Im(\Sigma_1)$ for every signature morphism $\sigma : \Sigma_1 \rightarrow \Sigma_2$. The right (entailment, closed) monotonic function commutes with the closure monotonic function: $\mathsf{right}_\Im^\vdash(\sigma) \cdot \mathsf{clo}_\Im(\Sigma_2) = \mathsf{clo}_\Im(\Sigma_1) \cdot \mathsf{right}_\Im^\bullet(\sigma)$ for every signature morphism $\sigma : \Sigma_1 \rightarrow \Sigma_2$. This means that the Galois connections commute $\phi_\Im^\bullet(\sigma) \circ \rho_\Im(\Sigma_1) = \rho_\Im(\Sigma_2) \circ \phi_\Im^\vdash(\sigma)$ for every signature morphism $\sigma : \Sigma_1 \rightarrow \Sigma_2$. For every signature Σ, closure (with a slight abuse of notation) forms a CPOG-morphism $\mathsf{clo}_\Im(\Sigma) = \langle 1_{|\mathsf{mod}|_\Im(\Sigma)}, \mathsf{incl}_\Im(\Sigma), \mathsf{clo}_\Im(\Sigma), 1_{\mathsf{sen}_\Im(\Sigma)}\rangle$: $\mathsf{cpog}_\Im(\Sigma) \rightleftharpoons \mathsf{clg}_\Im(\Sigma)$. For every signature morphism $\sigma : \Sigma_1 \rightarrow \Sigma_2$, we have the identity (commuting square) $\mathsf{clo}_\Im(\Sigma_1) \circ \mathsf{clg}_\Im(\sigma) = \mathsf{cpog}_\Im(\sigma) \circ \mathsf{clo}_\Im(\Sigma_2)$. Hence, closure is a natural transformation $\mathsf{clo}_\Im : \mathsf{cpog}_\Im \Rightarrow \mathsf{clg}_\Im \circ \mathsf{incl} : \mathsf{Sign}_\Im \rightarrow \mathsf{CPOG}$.

2.3   Theories

An institution $\Im$ has an existential quantification theory functor (indexed category) $\exists_\Im : (\mathsf{Sign}_\Im^{\mathsf{op}})^{\mathsf{op}} = \mathsf{Sign}_\Im \rightarrow \mathsf{Pre}$ to the category of (small) preorders and monotonic functions, where $\exists_\Im(\Sigma) = \mathsf{th}_\Im^\vdash(\Sigma)$ is the complete entailment preorder ("lattice of theories") for every signature Σ, and $\exists_\Im(\sigma) : \mathsf{th}_\Im^\vdash(\Sigma_1) \rightarrow \mathsf{th}_\Im^\vdash(\Sigma_2)$ is the direct existential image along the sentence function for every signature morphism $\sigma : \Sigma_1 \rightarrow \Sigma_2$. The Grothendieck construction on this functor produces a flattened category of theories $\mathsf{Theory}_\Im^\exists = \mathsf{Gr}(\exists_\Im)^{\mathsf{op}}$ (see the right side of Figure 1) whose objects are pairs $\langle \Sigma, T\rangle$ where T is a Σ-theory, and whose morphisms $\sigma : \langle \Sigma_1, T_1\rangle \rightarrow \langle \Sigma_2, T_2\rangle$ are signature morphsms $\sigma : \Sigma_1 \rightarrow \Sigma_2$ that map source axioms to target theorems $T_2^\bullet \supseteq \exists_\Im(\sigma)(T_1)$, or equivalently satisfy the constraint $\mathsf{sen}_\Im(\sigma)^{-1}(T_2^\bullet) \supseteq T_1$. There is an obvious base signature projection functor $\mathsf{base}_\Im^\exists : \mathsf{Theory}_\Im^\exists \rightarrow \mathsf{Sign}_\Im$. Composition and identities for theory

morphisms are defined in terms of their base signature morphisms. Based upon the limit theorem of the third section of the paper [12] on indexed categories and institutions, if $\mathsf{Sign}_\Im$ the category of signatures is cocomplete, then $\mathsf{Theory}_\Im^\exists$ the category of theories is cocomplete.

In the institutional approach of the IFF, unpopulated ontologies (no instance data and no classifications) are represented by theories, and the semantic integration of unpopulated ontologies is represented by the fusion (colimit construction) of theories. Fusion in the category of theories is direct existential image flow followed by meet in the lattices of theories (base signature fibers). The fusion construction is outlined in the following algorithm.

1. **Alignment:** Informally, identify the theories to be used in the construction. Decide on the semantic interconnection (semantic mapping) between theories. This may involve the introduction of some additional mediating (reference) theories.
   (a) Formally, create a diagram of theories $\mathcal{T}$ of shape (indexing) graph G that indicates this selection and interconnection. This diagram of theories is transient, since it will be used only for this computation. Other diagrams could be used for other fusion constructions.
   (b) Compute the base diagram of signatures $\mathcal{S} = \mathsf{base}_\Im^\exists(\mathcal{T})$ with the same shape. In more detail, $\mathcal{S} = \mathsf{base}_\Im^\exists(\mathcal{T}) = \{\mathcal{S}_k\} + \{\mathcal{S}_e : \mathcal{S}_m \to \mathcal{S}_n\} = \{\mathsf{base}_\Im^\exists(\mathcal{T}_k)\} + \{\mathsf{base}_\Im^\exists(\mathcal{T}_e) : \mathsf{base}_\Im^\exists(\mathcal{T}_m) \to \mathsf{base}_\Im^\exists(\mathcal{T}_n)\}$. Form the colimit (fusion) signature $\Sigma = \coprod \mathcal{S}$ of this diagram, with signature fusion cocone $\sigma : \mathcal{S} \Rightarrow \Sigma$. In more detail, $\sigma = \{\sigma_k : \mathcal{S}_k \to \Sigma\}$, satisfying the conditions $\sigma_m = \mathcal{S}_e \cdot \sigma_n$ for G-edge $e : m \to n$. Being the basis for theory colimits, signature colimits are important. They involve the two opposed processes of "summing" and "quotienting". Summing can be characterized as "keeping things apart" and "preserving distinctness", whereas quotienting can be characterized as "putting things together", "identification" and "synonymy". The "things" involved here are symbolic, and may involve relation type symbols, entity type symbols and the concepts that they denote.
2. **Unification:** Form the colimit (fusion) theory $T = \coprod \mathcal{T}$ of this diagram of theories, with theory fusion cocone $\tau : \mathcal{T} \Rightarrow T$. The fusion cocone is a universal corelation [4] that connects the individual theories in the diagram to the fusion theory. The fusion theory may be virtual.
   (a) Move (the individual theories $\{\mathcal{T}_k\}$ in) the diagram of theories $\mathcal{T}$ from the "lattice of theory diagrams" $\mathsf{th}_\Im^\vdash(\mathcal{S})$ along the signature morphisms in the signature fusion cocone $\sigma : \mathcal{S} \Rightarrow \Sigma$ to the lattice of theories $\mathsf{th}_\Im^\vdash(\Sigma)$ using the direct image function, getting the homogeneous diagram of theories $\exists_\Im(\sigma)(\mathcal{T})$ with the same shape G, where each theory $\exists_\Im(\sigma)(\mathcal{T})_k = \exists_\Im(\sigma_k)(\mathcal{T}_k)$ has the same base signature Σ (the meaning of homogeneous).
   (b) Compute the meet (union) of the diagram $\exists_\Im(\sigma)(\mathcal{T})$ within the "lattice of theories" $\mathsf{th}_\Im^\vdash(\Sigma)$, getting the fusion theory $T = \coprod \mathcal{T} = \bigwedge_{\Im,\Sigma} \exists_\Im(\sigma)(\mathcal{T}) = \bigcup \exists_\Im(\sigma)(\mathcal{T})$. The signature fusion cocone is the base of the theory fusion cocone: $\sigma = \mathsf{base}_\Im(\tau) : \mathsf{base}_\Im(\mathcal{T}) \Rightarrow \mathsf{base}_\Im(T)$.

3 Institutions in the Information Flow Framework

The main application of the IFF is institutional. To be independent of the logic used, the IFF represents and manipulates ontological structures within the metatheory of institutions. The IFF has work-in-progress axiomatizations for the institutions and connecting institution morphisms of information flow (**IF**), equational logic (**EQL**), order-sorted first order logic (**FOL**) and the common logic standard (**CL**), and is developing an axiomatization for the metatheory of institutions itself. In particular, the IFF has three work-in-progress efforts (meta-ontologies) that center around first order logic (FOL). Two of these efforts (`IFF-ONT` and `IFF-OO`) are non-traditional and the third effort (`IFF-FOL`) is traditional.

3.1 `IFF-ONT`

The IFF Ontology (meta) Ontology (`IFF-ONT`) is an older non-traditional axiomatization for FOL, It is based on the view that n-ary relations incorporate a notion of hypergraphs. The `IFF-ONT` has finished axiomatizations for the concepts (and categories) of signature (aka language), theory, model and (local) logic. The categories of signatures and theories are cocomplete. Free models and logics exist. `IFF-ONT` models are nonstandard: one component is a subset of abstract tuples (in the extreme, an `IFF-ONT` model might consist of just one abstract tuple), and thus are much more flexible than ordinary FOL models; and these models have a novel definition via classifications, $t \models r$ instead of $r(t)$. However, a problem was discovered during the axiomatization of the `IFF-ONT`. In order to represent an institution, model fiber (reduct) functors must be defined. For these definitions, the variable component of the object functions of these model functors is required to be a bijection. Hence, signatures and theories must be restricted to the corresponding subcategories. However, these subcategories are not cocomplete. But, completeness is a desirable property of the represented institution. The problem was isolated to the requirement that variables in the `IFF-ONT` have a fixed sort. This is apparently too inflexible, and in the newer effort, variables are sorted on the fly.

3.2 `IFF-OO`

The IFF Ontology (meta) Ontology (`IFF-OO`) is a newer non-traditional axiomatization for FOL. It is still based on view that n-ary relations incorporate a notion of hypergraphs. In fact, the `IFF-OO` identifies signatures (aka languages) with hypergraphs (of a certain kind). This kind of hypergraph is flexible, since it does not use a reference (sort) function. This means that node indices, which correspond to entity type variables, do not use a fixed sorting. Model fiber (reduct) functors exist, and the categories of signatures and theories are cocomplete. A special advantage has been found for this axiomatization: hypergraphs (signatures) are categorically equivalent to spangraphs. Spangraphs, which are a more flexible definition of signature in terms of relational arity, nicely model the

common logic standard (CL) "role-set syntax" notation, where arguments form a set of role-value pairs. For example, the sentence '`(Married (role-set: (wife Jill) (husband Jack)))`' expresses a marital relationship. Pat Hayes (see the CL archive) has advocated this notation in CL development. Position and argument order are not needed. According to Hayes, this provides some insurance against communication errors.

3.3   IFF-FOL

The IFF First Order Logic (meta) Ontology (`IFF-FOL`) gives a traditional axiomatization for FOL. The `IFF-FOL` has a very modular architecture (Figure 2), describing several institutions and institution morphisms along a spectrum, including **EQN** the institution for equational logic and **FOL** the institution for first order logic. Each edge in the diagram of (Figure 2) is associated with institution morphisms in both directions, projection downward and inclusion upward. The central bifurcation is between terms and expressions. FOL signatures are the pullback of expression signatures and term signatures over (bijections of) variables. FOL signatures with equality are the pullback of FOL signatures and universal algebra (equational signatures) over term signatures. Term signatures consist of function (type) symbols and variables. The Lawvere construction is defined here. Equations can be added giving equational signatures (equational presentations) as an extension of term signatures. They define a quotient of their Lawvere category. Expression signatures consist of relation (type) symbols and variables. Peircian existential graphs can be included here. FOL signatures consist of function (type) symbols, relation (type) symbols and variables. From the modular perspective of Figure 2, an FOL signature is a term signature and an expression signature that share a common set of variables.
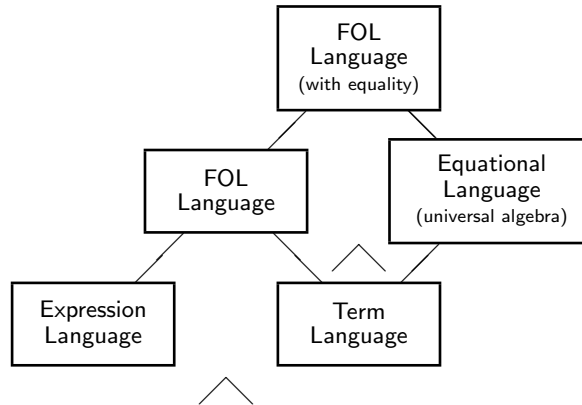

Fig. 2. The `IFF-FOL` module hierarchy

4   Summary

The IFF is a descriptive category metatheory whose architecture contains nested metalanguages. The main institutional application of the IFF is axiomatized in the Ontology (meta) Ontology (`IFF-ONT`, `IFF-OO`) and the FOL (meta) Ontology (`IFF-FOL`). These form flexible and general institutions for first order logic. Institutions formally express semantic integration as an ontological fusion process. Each community represents their conceptual space in their own terms, and connects with others via morphisms that enable ontological alignment specification. Institutions represent logical environments and institution morphisms connect logical environments. The IFF is in the process of axiomatizing the theory of institutions.

References

1. Kent, R.E.: The SUO Information Flow Framework (SUO IFF). Technical report, Institute of Electrical and Electronics Engineers (2001) Published as the webpage http://suo.ieee.org/IFF/].
2. Goguen, J.: Information integration in institutions. Draft paper for the Jon Barwise memorial volume edited by Larry Moss (2005)
3. Kent, R.E.: Semantic integration in the IFF. In Doan, A., Halevy, A., Noy, N., eds.: Semantic Integration 2003. Volume 82 of CEUR Workshop Proceedings., Sun SITE Central Europe (CEUR) (2003) Proceedings of the Semantic Integration Workshop at ISWC-03, Sanibel Island, Florida, USA, October 20, 2003.
4. Goguen, J.: Three perspectives on information integration. In Kalfoglou, Y., Schorlemmer, M., Sheth, A., Staab, S., Uschold, M., eds.: Semantic Interoperability and Integration. Number 04391 in Dagstuhl Seminar Proceedings, Dagstuhl Research Online Publication Server (2005)
5. Sowa, J.F.: Knowledge Representation: Logical, Philosophical, and Computational Foundations. Brookes/Coles (2000)
6. Lawvere, F.W., Rosebrugh, R.: Sets for Mathematics. Cambridge University Press (2003)
7. Mossakowski, T., Goguen, J., Diaconescu, R., Tarlecki, A.: What is a logic? In: Proceedings of the First World Conference on Universal Logic. (2005)
8. Lehmann, F., Wille, R.: A triadic approach to formal concept analysis. Lecture Notes in Artificial Intelligence **954** (1995) 32–43
9. Goguen, J., Burstall, R.: Institutions: Abstract model theory for specification and programming. Journal of the Association for Computing Machinery **39** (1992) 95–146 Preprint, Report CSLI-85-30, Center for the Study of Language and Information, Stanford University, 1985.
10. Barwise, J., Seligman, J.: Information Flow: The Logic of Distributed Systems. Volume 44 of Cambridge Tracts in Theoretical Computer Science. Cambridge University Press (1997)
11. Kent, R.E.: Distributed conceptual structures. In de Swart, H., ed.: Sixth International Workshop on Relational Methods in Computer Science. Volume 2561 of Lecture Notes in Computer Science., Springer (2002) 104–123
12. Tarlecki, A., Burstall, R., Goguen, J.: Some fundamental algebraic tools for the semantics of computation, part 3: Indexed categories. Theoretical Computer Science **91** (1991) 239–264